

2. 学会誌・雑誌等における論文掲載

掲載した論文（発表題目）	発表者氏名	発表した場所（学会誌・雑誌等名）	発表した時期	国内・外の別
再生医療等製品開発における動物実験—指針及び評価指標について—	澤田留美	オベリスク, 20(1), 25-31	2015	国内
再生医療等製品とバイオマテリアル, そして評価指標	澤田留美	バイオマテリアル—生体材料—, 33(1), 7-8	2015	国内
Label-free morphology-based prediction of multiple differentiation potentials of human mesenchymal stem cells for early evaluation of intact cells.	Sasaki H., Takeuchi I., Okada M., Sawada R., Kanie K., Kiyota Y., Honda H., and Kato R.	PLOS ONE, 9(4), e93952	2014	国外
Cyclin D2 promotes the proliferation of human mesenchymal stem cells.	Kono K., Niimi S., and Sawada R.	J. Bone Marrow Res., 2: 136. 1000136	2013	国外
Characterization of the cell growth analysis for detection of immortal cellular impurities in human mesenchymal stem cells.	Kono K, Takada N, Yasuda S, Sawada R, Niimi S, Matsuyama A, Sato Y.	Biologicals. 43:1 46-149	2015年	国外
The role of maternal-specific H3K9me3 modification in establishing imprinted X-chromosome inactivation and embryogenesis in mice.	Fukuda A, Tomikawa J, Miura T, Hata K, Nakabayashi K, Eggan K, Akutsu H, Umezawa A.	Nat Commun. 5:5464.	2014年	国外
Genome-wide copy number analysis and systematic mutation screening in 58 patients with hypogonadotropic hypogonadism.	Izumi Y, Suzuki E, Kanzaki S, Yatsuga S, Kinjo S, Igarashi M, Maruyama T, Sano S, Horikawa R, Sato N, Nakabayashi K, Hata K, Umezawa A, Ogata T, Yoshimura Y, Fukami M.	Fertil Steril. 102(4):1130-1136 .e3.	2014	国外

Compilation of copy number variants identified in phenotypically normal and parous Japanese women.	Migita O, Maehara K, Kamura H, Miyakoshi K, Tanaka M, Morokuma S, Fukushima K, Shimamoto T, Saito S, Sago H, Nishihama K, Abe K, Nakabayashi K, Umezawa A, Okamura K, Hata K.	J Hum Genet. 59 (6) :326-331.	2014	国外
Characterization of in vivo tumorigenicity tests using severe immunodeficient NOD/Shi-scid IL2Rgamma null mice for detection of tumorigenic cellular impurities in human cell-processed therapeutic products.	Kusakawa S, Machida K, Yasuda S, Takada N, Kuroda T, Sawada R, Okura H, Tsutsumi H, Kawamat H, Sato Y.	Regenerative Therapy. 1: 30-37	2015年	国内
A novel in vitro method for detecting undifferentiated human pluripotent stem cells as impurities in cell therapy products using a highly efficient culture system.	Tano K, Yasuda S, Kuroda T, Saito H, Umezawa A, Sato Y.	PLoS One. 9:e110496	2014年	国外
In vitro detection of residual undifferentiated cells in retinal pigment epithelial cells derived from human induced pluripotent stem cells.	Kuroda T, Yasuda S, Sato Y.	Methods in Molecular Biology 1210:183-92.	2014年	国外
再生医療と薬学	佐藤陽治	ファルマシア 50:1213-5.	2014年	国内
ヒト多能性幹細胞加工製品に残存する未分化多能性幹細胞の高感度検出法の開発	佐藤陽治	再生医療 13:432-5.	2014年	国内
再生医療・細胞治療の規制動向とレギュラトリーサイエンス	村岡ひとみ, 佐藤陽治	DDS 29:207-16.	2014年	国内
薬事法改正と再生医療等安全性確保法を踏まえた再生医療／細胞治療の開発	中島啓行, 佐藤陽治	ファームステージ 10:1-5.	2014年	国内
再生医療・細胞治療に使用する細胞加工物の品質・安全性評価の原則と造腫瘍性の考え方	三浦巧, 佐藤陽治	谷本学校毒性質問箱 16:1-10.	2014年	国内

再生医療製品の造腫瘍性評価	草川森士, 佐藤陽治	最新医学 69(3):745-765.	2014 年	国内
再生医療・細胞治療の臨床研究から実用化までの道のり	村岡ひとみ, 佐藤陽治	Geriatric Medicine (老年医学) 52(3):237-239.	2014 年	国内
ヒト iPS 細胞由来移植細胞中に混入する造腫瘍性細胞／未分化細胞の in vitro 検出法	佐藤陽治	Cytometry Research 24(1): 7-11.	2014 年	国内

再生医療等製品開発における動物実験

—指針及び評価指標について—

澤田 留美

国立医薬品食品衛生研究所／再生・細胞医療製品部

1. はじめに

わが国の再生医療分野における技術開発の進展は目覚しく、再生医療の実用化に向けた動きが急速に進められており、この分野での製品化を加速させるために現在、国家レベルでの取り組みがなされている。再生医療等製品とは、ヒトまたは動物の細胞に培養その他の加工を施したものであって身体の構造または機能の再建、修復または形成や、疾病の治療または予防を目的として使用されるものであり、製品の形態としては、細胞のみを用いたものだけでなくマトリックスやスキャホールド等の非細胞材料と細胞とのコンビネーション製品も想定される。

平成 25 年 11 月に「薬事法等の一部を改正する法律（平成 25 年法律第 84 号）」が公布され、平成 26 年 11 月にこれまでの「薬事法」が「医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律」と改められ、施行されることとなった。これまでの薬事法下での審査によりわが国で製造承認されている再生医療製品は 2 品目であり、薬事法下ではどちらも「医療機器」と分類されていた。しかし、薬事法の改正による医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律では、これまで「医薬品」もしくは「医療機器」の 2 つの分類であったものに加えて、「再生医療等製品」を新たに定義し、その特性を踏まえた安全性対策等の規制が設けられる事となった。

これまでに再生医療等製品の承認審査における品質や安全性評価等に関わる指針等としては、細胞・組織加工医薬品等の品質及び安全性の確保に関する指針が 2 つ（表 1）、幹細胞加工医薬品等の品質及び安全性の確保に関する指針が 5 つ（表 2）、さらに今後上市される可能性が高い再生医療等製品を想定し、それぞれの製品に特化した内容の評価指標が 7 つ（表 3）示されている。

表 1 細胞・組織加工医薬品等に関する指針

ヒト(自己)由来細胞・組織加工医薬品等の品質及び安全性の確保に関する指針	H20.2.8 薬食発第 0208003 号
ヒト(同種)由来細胞・組織加工医薬品等の品質及び安全性の確保に関する指針	H20.9.12 薬食発第 0912006 号

表 2 幹細胞加工医薬品等に関する指針

ヒト(自己)体性幹細胞加工医薬品等の品質及び安全性の確保に関する指針	H24.9.7 薬食発第 0907 第 2 号
ヒト(同種)体性幹細胞加工医薬品等の品質及び安全性の確保に関する指針	H24.9.7 薬食発第 0907 第 3 号
ヒト(自己)iPS(様)細胞加工医薬品等の品質及び安全性の確保に関する指針	H24.9.7 薬食発第 0907 第 4 号
ヒト(同種)iPS(様)細胞加工医薬品等の品質及び安全性の確保に関する指針	H24.9.7 薬食発第 0907 第 5 号
ヒトES細胞加工医薬品等の品質及び安全性の確保に関する指針	H24.9.7 薬食発第 0907 第 6 号

表 3 次世代医療機器評価指標

重症心不全細胞治療用細胞シートに関する評価指標	H22.1.18 薬食機発第 0118 第 1 号
角膜上皮細胞シートに関する評価指標	H22.1.18 薬食機発第 0118 第 1 号
角膜内皮細胞シートに関する評価指標	H22.5.28 薬食機発第 0528 第 1 号
関節軟骨再生に関する評価指標	H22.12.15 薬食機発第 1215 第 1 号
歯周組織治療用細胞シートに関する評価指標	H23.12.7 薬食機発第 1207 第 1 号
自己 iPS 細胞由来網膜色素上皮細胞に関する評価指標	H25.5.29 薬食機発第 0529 第 1 号
同種 iPS(様)細胞由来網膜色素上皮細胞に関する評価指標	H26.9.12 薬食機参発第 0912 第 2 号

本稿では、それぞれの指針及び評価指標に示されている動物実験の内容について紹介し、再生医療等製品開発における動物実験の位置づけについて考えてみたい。各指針では、「非臨床安全性試験」、「効力または性能を裏付ける試験」、「体内動態」の項目において、動物実験について言及されている。また、各評価指標においては、対象の製品に特化した内容の動物実験（疾患動物モデル）や造腫瘍性についても示されている。造腫瘍性試験のガイドラインとして、ICH-Q5D や WHO-TRS878 が参考とされる場合が多いが、WHO-TRS878 の 2010 年 7 月の改正案では、患者に移植する動物由来生細胞及び細胞・組織利用製品の原料となる細胞は「対象外」とされ、現段階では、再生医療製品を対象とした造腫瘍試験のガイドラインは実質存在しないと考えられる。それらをふまえた上で、造腫瘍試験における動物実験についても触れてみたい。

2. 細胞・組織加工医薬品等の品質及び安全性の確保に関する指針について

これまでに発出されている細胞・組織加工医薬品等の品質及び安全性の確保に関する指針は、①平成 20 年 2 月 8 日付薬食機発 0208003 号厚生労働省医薬食品局長通知「ヒト(自己)由来細胞・組織加工医薬品等の品質及び安全性の確保に関する指針」と②平成 20 年 9 月 12 日付薬食機発 09120006 号厚生労働省医薬食品局長通知「ヒト(同種)由来細胞・組織加工医薬品等の品質及び安全性の確保に関する指針」の二つである。両指針は、対象がそれぞれ使用する細胞・組織の由来が自己か同種かによって分かれるが、どちらも 7 章で構成され、第 1 章 総則、第 2 章 製造方法、第 3 章 細胞・組織加工医薬品等の安定性、第 4 章 細胞・組織加工医薬品等の非臨床安全性試験、第 5 章 細胞・組織加工医薬品等の効力又は性能を裏付ける試験、第 6 章 細胞・組織加工医薬品等の体内動態、第 7 章 臨床試験となっており、その内、第 4 章、第 5 章、第 6 章で動物実験について言及している。以下に、動物実験に関わる部分を抜粋して紹介する。

第 4 章 細胞・組織加工医薬品等の非臨床安全性試験では、「製品の特性及び適用法から評価が必要と考えられる安全性関連事項について技術的に可能であれ

ば、科学的合理性のある範囲で適切な動物を用いた試験または *in vitro* での試験を実施すること。」としている。さらに、その試験系の妥当性について明らかにした上で「動物由来の製品モデルを作成し適切な実験動物に適用する試験系により試験を行う」ことの科学的合理性の可能性も言及している。また、同種由来の製品については、非臨床的に安全性を確認する際に参考にすべき留意点の例の一つに、「株化細胞を用いた場合には、適切な動物モデル等を利用し、腫瘍形成及びがん化の可能性について考察し、明らかにすること。」と示している。

第 5 章 細胞・組織加工医薬品等の効力又は性能を裏付ける試験では、「実験動物又は細胞等を用い、細胞・組織加工医薬品等の機能発現、作用持続性及び医薬品・医療機器として期待される効果を検討すること。」としている。また、「疾患モデルがある場合には、それを用いて治療効果を検討すること。」とも示されている。

第 6 章 細胞・組織加工医薬品等の体内動態では、「実験動物での吸収及び分布等の体内動態に関する試験等により、患者等に適用された製品中の細胞・組織の生存期間、効果持続期間を推察し、目的とする効果が十分得られることを明らかにすること。」とし、「当該細胞・組織が特定の部位(組織等)に到達して作用する場合には、その局在性を明らかにすること。」としている。

なお、第 4 章のみならず、第 5、6 章とも、上記の内容の前に「技術的に可能かつ科学的に合理性のある範囲で」と示されていることも付記しておく。

3. 幹細胞加工医薬品等の品質及び安全性の確保に関する指針について

2. に示した平成 20 年に発出された 2 指針が対象にしている製品の由来の「細胞」のうち「幹細胞」を加工した医薬品又は医療機器の品質及び安全性の確保のための基本的な技術要件については、平成 24 年に新たな指針が発出された。それは、①平成 24 年 9 月 7 日付薬食機発第 0907 第 2 号厚生労働省医薬食品局長通知「ヒト(自己)体性幹細胞加工医薬品等の品質及び安全性の確保に関する指針」、②同第 3 号「ヒト(同種)体性幹細胞加工医薬品等の品質及び安全性の確保に関する指針」、③同第 4 号「ヒト(自己)iPS(様)細胞加工医

薬品等の品質及び安全性の確保に関する指針」、④同第5号「ヒト(同種)iPS(様)細胞加工医薬品等の品質及び安全性の確保に関する指針」、⑤同第6号「ヒトES細胞加工医薬品等の品質及び安全性の確保に関する指針」の5指針である。これらの5指針は、対象がそれぞれ使用する幹細胞の種類と自己/同種によって分かれるが、その構成は、2.の細胞・組織加工医薬品等の品質及び安全性の確保に関する指針と同様であり、第4章、第5章、第6章にて動物実験について言及している。以下に、動物実験に関わる部分を抜粋して紹介する。

第4章の非臨床安全性試験では、平成20年の2指針と同様に「製品の特性及び適用法から評価が必要と考えられる安全性関連事項について技術的に可能であれば、科学的合理性のある範囲で適切な動物を用いた試験又は *in vitro* での試験を実施すること。」とし、また「動物由来の製品モデルを作成し適切な実験動物に適用する試験系により試験を行う」ことの科学的合理性の可能性も言及している。さらに、2.の細胞・組織加工医薬品等の指針では特に記述はなかったが、幹細胞加工医薬品等の指針においては「対象疾患ごとに適切なモデル動物を用いた試験の実施を考慮する」とし、「例えば神経疾患ならばサル等、循環器疾患ならばブタ・イヌ等が適している場合がある」と具体的に動物種を例示している。また、非臨床的に安全性を確認する際に参考にすべき留意点の例の一つに、「腫瘍形成及びがん化の可能性について」の事項があり、「必要に応じて適切な動物実験モデル等を利用した検討を行うこと。」とし、「動物モデルを使用した造腫瘍性試験においては、細胞の分散や足場への接着、細胞密度、投与部位等の条件が最終製品と必ずしも一致するものではない。さらに、動物の種・系統・免疫状態による感度差もある。これらの事情を総合的に勘案して、最終製品の造腫瘍性を評価する必要がある。」としている。また、③、④、⑤のiPS(様)細胞およびES細胞加工医薬品等の非臨床安全性試験では、「最終製品における未分化細胞の存在が異所性組織形成や腫瘍形成・がん化の可能性など安全性上の重要な関心事であるが、可能な限り、セル・バンクや中間製品段階等での徹底的な解析により、混在の可能性を否定するか、あるいは、目的細胞から未分化細胞の効果的分離・除去法や不活化法を開発し、活用することにより、混在の可能性を最小限にする

努力が求められる。さらに、投与経路等の選択も安全性上の懸念を最小限にするための有用な方策である可能性がある。」と示されており、製品中の未分化細胞の存在についての注意喚起とともに投与経路等の選択の際の動物実験の寄与の可能性が伺われる。

第5章の効力又は性能を裏付ける試験及び第6章の体内動態については、平成20年の2指針(細胞・組織加工医薬品等の指針)と同様である。

4. 次世代医療機器・再生医療等製品評価指標について

次世代医療機器評価指標作成事業は、平成17年度より開始され、医療機器の開発の迅速化および薬事法審査の円滑化に資する評価指標の検討、作成を行うため、厚生労働省(審査WG担当)と経済産業省(開発WG担当)が合同で検討会を設置した。厚生労働省では、革新的な医療機器の開発の迅速化および薬事法審査の円滑化を図るため、近い将来上市される可能性が高い次世代医療機器の分野について、効率的な開発および迅速な審査が必要とされる技術要素を選定して、開発段階から承認審査を見通した審査に係る評価指標を検討、作成することとしたものである。平成26年10月までに発出された評価指標は、表4に示す。

表4 これまでに発出された次世代医療機器・再生医療等製品評価指標

対象分野	次世代医療機器・再生医療等製品評価指標	通知
再生医療	重症心不全細胞治療用細胞シートに関する評価指標	H22.1.18 薬食機発第0118第1号
	角膜上皮細胞シートに関する評価指標	H22.1.18 薬食機発第0118第1号
	角膜内皮細胞シートに関する評価指標	H22.5.29 薬食機発第0229第1号
	関節軟骨再生シートに関する評価指標	H23.12.7 薬食機発第0237第1号
	歯周組織再生用細胞シートに関する評価指標	H29.12.15 薬食機発第0215第1号
	自己iPS細胞由来網膜色素上皮細胞に関する評価指標	H25.5.29 薬食機発第0229第1号
	同種iPS(様)細胞由来網膜色素上皮細胞に関する評価指標	H26.9.12 薬食機発第0212第2号
体内埋め込み型材料	整形外科用骨接合材料カスチンインプラントに関する評価指標	H22.12.15 薬食機発第0215第1号
	整形外科用カスチンインプラントに関する評価指標	H23.12.7 薬食機発第0207第1号
	整形外科用カスチンインプラントに関する評価指標	H24.11.20 薬食機発第0120第3号
	可動性及び安定性を維持する骨接合インプラントに関する評価指標	H25.9.12 薬食機発第0212第2号
体内埋め込み型診断装置	二次元画像診断装置に関する評価指標	H23.4.4 薬食機発第0404第2号
	神経機能診断装置に関する評価指標	H22.12.15 薬食機発第0215第1号
ナビゲーション医療	骨析像診断装置に関する評価指標	H22.1.18 薬食機発第0118第1号
	関節手術支援装置に関する評価指標	H22.1.18 薬食機発第0118第1号
	軟組織に適用するコンピュータ支援手術装置に関する評価指標	H22.9.19 薬食機発第0219第1号
	コンピュータ支援手術装置に関する評価指標	H23.12.7 薬食機発第0207第1号
体外診断装置	DNAチップを用いた遺伝子型診断装置に関する評価指標	H24.4.4 薬食機発第0404第2号
	RNAプロファイルに基づいた診断装置に関する評価指標	H24.11.20 薬食機発第0120第3号
活動細胞回収装置	活動細胞回収装置に関する評価指標	H25.5.29 薬食機発第0229第1号
骨髄下貯血液作製	骨髄下貯血液作製用医療機器に関する評価指標	H25.5.29 薬食機発第0229第1号

本事業における審査WGの活動成果や評価指標の内容については、<http://dmd.nihs.go.jp/jisedai/>をご参照頂きたい。

再生医療分野の審査WGについてのこれまでの成果としては、7つの評価指標素案(①重症心不全細胞治療用細胞シート、②角膜上皮細胞シート、③角膜内皮細胞シート、④関節軟骨再生、⑤歯周組織治療用細胞シート、⑥自己iPS細胞由来網膜色素上皮細胞、⑦同種iPS(様)細胞由来網膜色素上皮細胞)を作成した。それらは、それぞれ厚生労働省医療機器審査管理室長通知(平成22年1月18日付薬食機発0118第1号別添3及び別添4、平成22年5月28日付薬食機発0528第1号別添1、平成22年12月15日付薬食機発1215第1号別添1、平成23年12月7日付薬食機発1207第1号別添1、平成25年5月29日付薬食機発0529第1号別添1)及び厚生労働省大臣官房参事官(医療機器・再生医療等製品審査管理室担当)通知(平成26年9月12日付薬食機参発0912第2号別紙1)として発出されている(表3)。これらは、すべて2. および3. で紹介した指針をベースとしてそれぞれの製品に特有の留意すべき事項を示している。以下にそれぞれの評価指標の内容について、動物実験について言及している非臨床試験の項目を中心に紹介していく。

①重症心不全細胞治療用細胞シートに関する評価指標(平成22年1月18日付薬食機発0118第1号別添3)

本評価指標は、2. で示したヒト(自己/同種)由来細胞・組織加工医薬品等の指針をベースにしたもので、その旨を「1. はじめに」の部分で明記している。その上で、「本評価指標は、ヒト由来細胞・組織加工医薬品等のうち特に重症心不全等の治療を目的として心臓に適用されるものであって細胞シート状の製品について、上述の基本的な技術要件のうち当該製品に特有の留意すべき事項を示すものである。」と定めている。動物実験については、5. 評価に当たって留意すべき事項の(2)非臨床試験にて言及している。「動物に細胞シート製品を適用して有効性及び安全性を評価する際には、対象疾患を考慮して疾患モデル動物を作成すること。」とし、さらに「疾患モデル動物作成方法の例として、冠動脈結紮による心筋梗塞モデル又は overdriving 法による拡張型心筋症モデル等が挙げられるが、用いた動物モデ

ルについては、その選択の根拠、試験系の妥当性及び得られた結果のヒトへの外挿性について説明する必要がある。有効性及び安全性の評価のために、細胞シート製品を適用した群、対照物質を適用した対照群、さらに必要であれば sham オペ群を用いた比較試験を実施すること。評価期間についても、その設定根拠について説明すること。移植した細胞シート製品とそれによってもたらされる効能について、細胞シートに含まれる細胞の移植部位における局在性を確認するなど経時的に評価し、その関連性について考察すること。また、動物試験は、使用方法に関する試験の意味合いも含んでいることから、動物への適用方法は、可能な限り臨床での使用法(例えば、開胸手術、内視鏡的手術など)を反映することが望ましい。」とかなり具体的に疾患モデルや実験内容について示している。また、評価項目の一つとして、催不整脈性の評価が挙げられ、「催不整脈性に関しては、普遍的に受容されたモデル動物は確立されていないため、対象疾患等を考慮して作成した疾患モデル動物(イヌまたはブタなど)を用いて評価すること。」としている。

②角膜上皮細胞シートに関する評価指標(平成22年1月18日付薬食機発0118第1号別添4)

本評価指標は、①と同様にヒト(自己/同種)由来細胞・組織加工医薬品等の指針をベースにしたもので、「本評価指標は、ヒト由来細胞・組織加工医薬品等のうち特に角膜上皮障害等の治療を目的として適用されるものであって細胞シート状の製品について、上述の基本的な技術要件のうち当該製品特有の留意すべき事項を示すものである。」と定めている。動物実験に関しては、5. (2)非臨床試験の中で「妥当な疾患動物モデルと、そのモデル動物に由来する原材料を用いた同等品作製および品質保証技術が確立されている場合には、可能な限り、モデル動物を用いた非臨床試験を行うことが望ましい。」としている。さらに「疾患モデル動物としては、例えば、ウサギ、イヌ、ブタ、サル等の中大型動物を用いた角膜上皮幹細胞疲弊症モデル等が考えられる。」とモデル動物を具体的に例示している。また「用いた動物モデルの作製方法、試験デザイン、評価期間、評価方法等の妥当性について説明する必要がある。有効性の検討には、適切な対照群を設定し、統計学的手法を用いて評価することが望ましい。」とも示されている。

検査項目としては、「①細隙灯顕微鏡検査による角膜の評価や②形態学的評価」が考えられるとし、その評価法等も示されている。

③角膜内皮細胞シートに関する評価指標(平成22年5月28日付薬食機発0528第1号別添1)

本評価指標は、①と同様にヒト(自己/同種)由来細胞・組織加工医薬品等の指針をベースにしたもので、「本評価指標は、ヒト由来細胞・組織加工医薬品等のうち特に角膜内皮障害等の治療を目的として適用されるものであって細胞シート状の製品について、上述の基本的な技術要件に加えて当該製品特有の留意すべき事項を示すものである。」と定めている。動物実験に関しては、5.(2)非臨床試験の中で「妥当な疾患動物モデルと、そのモデル動物に由来する原材料を用いた同等品作製および品質保証技術が確立されている場合には、可能な限り、モデル動物を用いた非臨床試験を行うことが望ましい。」とし、「疾患モデル動物としては、例えば、中大型動物を用いた角膜内皮機能不全モデル等が考えられる。」としている。検査項目としては、「①細隙灯顕微鏡検査による角膜の評価、②角膜厚の評価、③スペキュラーマイクロスコープによる角膜内皮細胞の密度・形態の評価、④形態学的評価、⑤安全性の評価」を示しており、それぞれの評価法についても言及している。

④関節軟骨再生に関する評価指標(平成22年12月15日付薬食機発1215第1号別添1)

本評価指標もやはり、①と同様にヒト(自己/同種)由来細胞・組織加工医薬品等の指針をベースにしたもので、「本評価指標は、ヒト由来細胞・組織加工医薬品等のうち特に損傷関節軟骨等の治療を目的として軟骨に適用される、ヒト軟骨細胞加工医薬品若しくは医療機器(以下「ヒト軟骨細胞加工医薬品等」という。)又はヒト間葉系幹細胞加工医薬品若しくは医療機器(以下「ヒト間葉系幹細胞加工医薬品等」という。)について、上述の基本的な技術要件に加えて留意すべき事項を示すものである。」と定めている。動物実験に関しては、まず5.評価に当たって留意すべき事項の(9)細胞の造腫瘍性・過形成において、「悪性腫瘍のみならず、良性腫瘍を含む腫瘍形成及び過形成の可能性を検討すること。」とし、「試験により造腫瘍性を評価する方法としては、例えば核型分析、軟寒天コロニー形成試験、免疫

不全動物における腫瘍形成能試験等が挙げられる。」と動物実験を例示している。また、「なお、免疫不全動物における腫瘍形成能試験においては、移植した細胞が体内で軟骨を形成した場合も腫瘍のように見えることがあるので、形態的特徴だけでなく組織病理学的特徴による評価も検討すること。」と関節軟骨再生に特化した留意点も挙げられている。さらに、「免疫不全動物における腫瘍形成能試験については、WHO Expert Committee on Biological Standardization. Forty-Seventh Report(1998)等を参考にすることが考えられるが、試験法の妥当性については、製品の特性やその時点での技術レベル等に応じて検討を行うこと。」とも記されている。また、6.効力又は性能を裏付ける試験については、「適当な動物由来細胞・組織製品モデル又は関節疾患モデルがある場合には、それを用いて治療効果を検討すること。」とし、治療効果の評価方法についても言及している。さらに、7.体内動態については、「技術的に可能で科学的合理性のある範囲で、実験動物での分布、吸収、遊走、生着等の体内動態に関する試験を実施すること。試験を実施しない場合には、その妥当性を示すこと。」としている。

⑤歯周組織治療用細胞シートに関する評価指標(平成23年12月7日付薬食機発1207第1号別添1)

本評価指標もやはり、①と同様にヒト(自己/同種)由来細胞・組織加工医薬品等の指針をベースにしたもので、「本評価指標は、歯周組織破壊を伴う歯周疾患(歯周炎)等の治療を目的として適用されるヒト骨膜細胞加工医療機器、ヒト歯根膜細胞加工医療機器、ヒト骨髄由来間葉系幹細胞加工医療機器及びヒト脂肪由来間葉系幹細胞加工医療機器のうち細胞シート状の製品(支持体が含有された製品を含む。以下同じ。)について、基本的な技術要件に加えて品質、有効性及び安全性の評価にあたって留意すべき事項を示すものである。」と定めている。動物実験に関しては、5.(2)非臨床試験の中で、「非臨床試験としては主に大型動物(ブタ・サルが用いられる場合もあるが主にイヌ)が用いられる。欠損モデルとしては骨内欠損モデル(1壁性、2壁性、3壁性)、分岐部モデル(Ⅱ度、Ⅲ度)、水平性クリティカルディフェクトモデルなどの様々な欠損モデルが用いられている。実験的欠損を切削器具などで作製し、その際に炎症を惹起させるマテリアルを欠損部に充填させ

ることにより意図的な炎症モデルを作成することも可能である。」と動物モデル等が具体的に示されており、「少なくとも上記のいずれかのモデルにおいて歯周組織治療用細胞シートの有効性を示す必要がある。」としている。さらに、「有効性の評価はX線的ならびに組織学的に、歯槽骨・歯根膜・セメント質の形態計測等により行う。」と有効性評価についても言及している。

⑥自己 iPS 細胞由来網膜色素上皮細胞に関する評価指標(平成 25 年 5 月 29 日付薬食機発 0529 第 1 号別添 1)

平成 26 年 9 月に、理化学研究所の高橋政代プロジェクトリーダーらのチームにより、加齢黄斑変性の患者に、患者の iPS 細胞から作った網膜色素上皮細胞を移植する世界初の手術が行われた。本評価指標はそれに先駆けて作成され、まさに本事業の目的である「近い将来上市される可能性が高い分野について、効率的な開発および迅速な審査が必要とされる技術要素を選定して、開発段階から承認審査を見通した審査に係る評価指標を検討、作成」したものである。

本評価指標は、平成 24 年 9 月 7 日付薬食機発第 0907 第 4 号厚生労働省医薬食品局長通知「ヒト(自己) iPS(様)細胞加工医薬品等の品質及び安全性の確保に関する指針」をベースにしたもので、「本評価指標は、ヒト(自己) iPS 細胞加工医薬品等のうち特に網膜色素上皮障害等の治療を目的として適用される医療機器について、上述の基本的な技術要件に加えて当該製品特有の留意すべき事項を示すものである。」と定めている。動物実験に関しては、5. (2)非臨床試験の中の 3 つの項目で言及している。まず、①造腫瘍性試験では、「製造工程の妥当性の検証として、同じ方法で作製した同じ品質基準を満たす最終製品について可能であれば別個体より得られた 3 ライン以上を用い、免疫不全動物を利用した検出感度が既知の試験系を用いて一定数以上の局所(網膜下あるいは皮下)における造腫瘍性の評価を行うことが有用である。」としているが、「ただし、これらの造腫瘍性の検討に関しては、今後の臨床応用でのデータの蓄積に応じ見直しが必要となる可能性もある。」とも記されている。さらに、「なお、造腫瘍性試験を含む非臨床安全性に関する一般的な考え方として、最終製品の安全性と原材料(iPS 細胞)の安全性は原則区別されるものである。」としている。次に、②効力又

は性能を裏付ける試験では、「網膜色素上皮移植については、ほぼ機能的に成熟した状態の細胞を移植することになるので、原則、RCS ラット(Royal College of Surgeons rat)など網膜色素上皮機能不全動物モデル網膜下に移植を行い、網膜色素上皮細胞としての網膜保護効果を確認する。」と具体的な動物モデルも挙げられている。そして③その他では、「シート挿入などで特別な手技を必要とする場合、手技的な安全性の確認、及びその手技を用いての移植後の局所における短期間での反応など、臨床応用において必要かつ科学的に妥当と思われる項目については、中、大型動物での確認を行う事が望ましい。」としている。

⑦同種 iPS(様)細胞由来網膜色素上皮細胞に関する評価指標(平成 26 年 9 月 12 日付薬食機参発 0912 第 2 号別紙 1)

本評価指標は、iPS 細胞由来網膜色素上皮細胞を製品化するためにはその材料として欠かすことのできない同種由来の細胞を対象とした。ベースとなるのは、平成 24 年 9 月 7 日付薬食機発第 0907 第 5 号厚生労働省医薬食品局長通知「ヒト(同種) iPS(様)細胞加工医薬品等の品質及び安全性の確保に関する指針」である。動物実験に関しては、5. (4)の、①最終製品の品質管理又は非臨床安全性評価のための造腫瘍性試験では、「造腫瘍性試験については最終製品を用い、免疫不全動物を利用した検出限界が既知の試験系を用いて造腫瘍性の評価を行うことが有用である。」とし、「品質管理」のための造腫瘍性試験(造腫瘍性細胞の存在量の確認)の例としては観察の簡便性と高感度な特性から、重度免疫不全動物(例: NOD/SCID/ γ Cnull (NOG) マウス、NOD/SCID/IL2r γ KO (NSG) マウス、Rag2- γ C double-knockout (DKO) マウス)への皮下投与試験が挙げられ、「非臨床安全性評価」のための造腫瘍性試験(最終製品の細胞がヒトでの投与部位に相当する微小環境で造腫瘍性を示すかどうかの確認)の例としては免疫不全動物への網膜下投与が挙げられる。」と、目的の違いによる造腫瘍性試験の考え方やそれぞれの実験方法等が例示されている。また、「網膜下(臨床投与経路)移植については、小動物では手術侵襲が大きく、手術手技により結果判定が困難となる可能性があることに留意する。」、さらに「網膜下移植による造腫瘍性試験の目的は、最終製品の細胞がヒトでの投与部位

に相当する微小環境で造腫瘍性を示すかどうかの確認にあることに留意しながら投与細胞数を設定することが重要である。」、そして「各セル・バンクから製造された最終製品について、ヒトでの投与部位に相当する微小環境で造腫瘍性を示すかどうかを評価する必要がある。免疫不全動物の網膜下への移植による最終製品の造腫瘍性試験は、その代表的な方法として挙げられる。」とし、iPS 細胞由来製品であるが故の造腫瘍性試験の重要性、同種細胞を用いるための品質管理、網膜色素上皮細胞であるための使用部位の限定等を鑑みた動物実験について言及している。また、②最終製品の効力または性能を裏付ける試験では、「技術的に可能かつ科学的に合理性のある範囲で、対象疾患に対し適切なモデル動物等を用いて、最終製品の機能発現、作用持続性、ヒト(同種)iPS(様)細胞加工製品として期待される臨床効果の実現可能性(Proof-of-Concept, POC)を示すこと。」としている。さらに、③その他では、「移植時の手技的な安全性の確認、その手技を用いての移植後の局所における短期間での反応等、臨床応用において必要かつ科学的に妥当と考えられる項目については、目的に応じて例えば中型又は大型動物を利用することにより確認を行うことが望ましい。」としている。

5. おわりに

平成 25 年 4 月に再生医療推進法が成立し、その中

に国の責務として再生医療の普及の促進が明記されている。再生医療分野の技術は日進月歩であり、特に再生医療への応用が期待されている iPS 細胞を用いた技術革新は目覚ましいものがある。前述したように、平成 26 年 9 月には iPS 細胞を用いた加齢黄斑変性の治療の臨床研究が始った。引き続き京都大学ではパーキンソン病、慶應義塾大学では脊髄損傷、大阪大学では心臓病の治療に iPS 細胞を用いた臨床研究が計画されている。さらに、京都大学 iPS 細胞研究所(CiRA)では、iPS 細胞を用いた再生医療等製品の実用化のために、世界標準として認知される高品質で安全な iPS 細胞の樹立法、維持培養法、品質評価法を確立し、数年内の iPS 細胞のバンク化をめざしている。自己だけでなく同種の細胞由来の高品質な iPS 細胞が安定的に供給されるようになれば、再生医療の実用化がより現実味を帯びる。この様に再生医療分野では、今後新しい技術の開発が加速的に進むことが予想され、それに伴い多種多様な(対象疾患、製品の形態など)再生医療等製品の開発に向けた臨床研究への橋渡しとなる非臨床試験における安全性および有効性の評価のための動物実験は、それぞれの対象製品に適した実験動物や手法が求められ、より重要な位置を占めていくことは間違いないであろう。

再生医療等製品とバイオマテリアル,そして評価指標

JJSB

バイオマテリアル-生体材料 33巻1号 平成27年

平成25年11月27日に「薬事法等の一部を改正する法律(平成25年法律第84号)」が公布され、平成26年11月25日にこれまでの「薬事法」が「医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律」と改められ、施行された。これまでの薬事法下での審査によりわが国で製造承認されている再生医療製品は2品目であり、薬事法下ではどちらも「医療機器」と分類されていた。しかし、医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律では、これまで「医薬品」もしくは「医療機器」の分類であったものに加えて、「再生医療等製品」を新たに定義し、その特性を踏まえた安全性対策等の規制が設けられることとなった。再生医療等製品とは、人又は動物の細胞に培養その他の加工を施したものであって人又は動物の、① 身体の構造又は機能の再建、修復又は形成、② 疾病の治療又は予防を目的として使用されるものである、と定義されている。製品の形態としては、細胞のみを用いたものだけでなく種々のマテリアルやスキャホールド等の非細胞材料と細胞や組織との複合製品も想定される。

ところで、「バイオマテリアル」とは一体何を指すのか、みなさんは正確にご存じでしょうか。“日本バイオマテリアル学会のご案内”によると「生体やその構成要素と直接あるいは間接に接触させて、傷んだ組織や器官、あるいは機能の診断や治療を行い、さらにある場合には損傷

提言



澤田留美

Rumi Sawada
国立医薬品食品衛生研究所
再生・細胞医療製品部第二室室長

部を補ったり、置き換えたりするために用いる材料」をバイオマテリアルとよぶらしい。そして、対象とする材料は、「金属、セラミックス、高分子、生体構成物質(糖類、タンパク質、DNA、細胞など)及びそれらの複合あるいはハイブリッド材料」と記されている。ということは、バイオマテリアルのうち、細胞を加工したもの(種々の材料との複合品も含む)が“再生医療等製品”となり、細胞を用いない金属やセラミックス等の材料は“医療機器”、加工していない細胞やその他の生体構成物質は“医薬品”と分類されることになる。つまり、傷んだ組織等の“再生”を目的とした治療に用いるバイオマテリアルであっても、細胞を加工しないものは再生医療等製品にはならない。再生医療等製品とバイオマテリアル

の関係は実はなかなか複雑である。

再生医療等製品の承認審査における品質や安全性評価等に関わる指針等としては、これまでに細胞・組織加工医薬品等の品質及び安全性の確保に関する指針が二つ(H20.2.8 薬食発第0208003号, H20.9.12 薬食発第0912006号)、幹細胞加工医薬品等の品質及び安全性の確保に関する指針が五つ(H24.9.7 薬食発第0907第2~6号)発出されている。さらに今後上市される可能性が高い再生医療等製品を想定しそれぞれの製品に特化した内容の評価指標が七つ示されているが、これらは次世代医療機器評価指標作成事業によって作成されている。本事業は平成17年度より開始され、医療機器の開発の迅速化および薬事法(当時)審査の円滑化に資する評価指標の検討、作成を行うために厚生労働省(審査ワーキンググループ(WG)担当)と経済産業省(開発WG担当)が合同で検討会を設置した。厚生労働省では、革新的な医療機器の開発の迅速化及び薬事法審査の円滑化を図るため、近い将来上市される可能性が高い次世代医療機器の分野について、効率的な開発および迅速な審査が必要とされる技術要素を選定して、開発段階から承認審査を見通した審査に係る評価指標を検討、作成することとしたものである。

筆者は、本事業の開始当時から現在に至るまで再生医療分野の審査WGの事務局として評価指標の作成のお手伝いをさせていただいている。

バイオマテリアル-生体材料 33-1, 2015 7

WGでは、臨床系、工学系、レギュラトリーサイエンスそれぞれの第一線で活躍されている専門家の先生方に委員としてお集まりいただき、国内外における最新の情報からそれぞれの対象製品に特化した品質、有効性及び安全性に関する評価指標を作成していただいているため、筆者はありがたいことにさまざまなバイオマテリアル開発関連を含む最新の研究に触れることができる貴重な機会を得ている。

本WGにて作成した評価指標は、①重症心不全細胞治療用細胞シート、②角膜上皮細胞シート、③角膜内皮細胞シート、④関節軟骨再生、⑤歯周組織治療用細胞シート、⑥自己iPS細胞由来網膜色素上皮細胞、⑦同種iPS(様)細胞由来網膜色素上皮細胞に関しての七つであり、それぞれ厚生労働省医療機器審査管理室長通知(①②H22.1.18 薬食機発0118第1号別添3, 4, ③H22.5.28 薬食機発0528第1号別添1, ④H22.12.15 薬食機発1215第1号別添

1, ⑤H23.12.7 薬食機発1207第1号別添1, ⑥H25.5.29 薬食機発0529第1号別添1)及び厚生労働省大臣官房参事官(医療機器・再生医療等製品審査管理室担当)通知(⑦H26.9.12 薬食機参発0912第2号別紙1)として発出されている。

自家培養軟骨(J-TEC)は平成24年に製造販売承認を取得したが、それに先駆け④関節軟骨再生に関する評価指標を作成している。また平成26年9月に理化学研究所の高橋政代プロジェクトリーダーらのチームにより加齢黄斑変性の患者に患者のiPS細胞からつくった網膜色素上皮細胞を移植する世界初の手術が行われ本格的な臨床研究がスタートしたが、やはりそれに先駆け⑥自己及び⑦同種iPS(様)細胞由来網膜色素上皮細胞に関する評価指標を作成している。このように本事業で作成された評価指標が実際の製品開発や製造承認申請に少しずつ関わることができはじめています。バイオマテリアルをご研究されている先生方が本稿で紹介さ

せていただいた評価指標について少しでも関心を持っていただき、先生方の研究開発時の一助となれば幸いです。評価指標の内容についてご興味のある方は<http://dmd.nihs.go.jp/jisedai/>をご参照願いたい。

平成25年4月に成立した再生医療推進法には、国の責務として再生医療の普及の促進が明記されている。再生医療分野の技術は日進月歩であり、今後新しい技術の開発が加速的に進むことが予想される。それに伴い、さまざまなバイオマテリアルの再生医療への応用が期待される。今後のバイオマテリアル研究開発の際の道標の一つとして評価指標を活用していただけたら何よりである。

末筆ではありますが、筆者のような若輩者に本稿の執筆の機会を与えてくださった、本誌編集委員長であり、かつまた評価指標作成事業においても大変お世話になりました東京理科大学教授菊池明彦先生に深謝致します。



Label-Free Morphology-Based Prediction of Multiple Differentiation Potentials of Human Mesenchymal Stem Cells for Early Evaluation of Intact Cells

Hiroto Sasaki¹, Ichiro Takeuchi², Mai Okada³, Rumi Sawada⁴, Kei Kanie^{1,3}, Yasujiro Kiyota⁵, Hiroyuki Honda¹, Ryuji Kato^{1,3*}

1 Department of Biotechnology, Graduate School of Engineering, Nagoya University, Nagoya, Aichi, Japan, **2** Department of Computer Science/Scientific and Engineering Simulation, Graduate School of Engineering, Nagoya Institute of Technology, Nagoya, Aichi, Japan, **3** Department of Basic Medicinal Sciences, Graduate School of Pharmaceutical Sciences, Nagoya University, Nagoya, Aichi, Japan, **4** Division of Medical Devices, National Institute of Health Sciences, Setagaya-ku, Tokyo, Japan, **5** Nikon Corporation, Chiyoda-ku, Tokyo, Japan

Abstract

Precise quantification of cellular potential of stem cells, such as human bone marrow-derived mesenchymal stem cells (hBMSCs), is important for achieving stable and effective outcomes in clinical stem cell therapy. Here, we report a method for image-based prediction of the multiple differentiation potentials of hBMSCs. This method has four major advantages: (1) the cells used for potential prediction are fully intact, and therefore directly usable for clinical applications; (2) predictions of potentials are generated before differentiation cultures are initiated; (3) prediction of multiple potentials can be provided simultaneously for each sample; and (4) predictions of potentials yield quantitative values that correlate strongly with the experimental data. Our results show that the collapse of hBMSC differentiation potentials, triggered by *in vitro* expansion, can be quantitatively predicted far in advance by predicting multiple potentials, multi-lineage differentiation potentials (osteogenic, adipogenic, and chondrogenic) and population doubling potential using morphological features apparent during the first 4 days of expansion culture. In order to understand how such morphological features can be effective for advance predictions, we measured gene-expression profiles of the same early undifferentiated cells. Both senescence-related genes (p16 and p21) and cytoskeleton-related genes (*PTK2*, *CD146*, and *CD49*) already correlated to the decrease of potentials at this stage. To objectively compare the performance of morphology and gene expression for such early prediction, we tested a range of models using various combinations of features. Such comparison of predictive performances revealed that morphological features performed better overall than gene-expression profiles, balancing the predictive accuracy with the effort required for model construction. This benchmark list of various prediction models not only identifies the best morphological feature conversion method for objective potential prediction, but should also allow clinicians to choose the most practical morphology-based prediction method for their own purposes.

Citation: Sasaki H, Takeuchi I, Okada M, Sawada R, Kanie K, et al. (2014) Label-Free Morphology-Based Prediction of Multiple Differentiation Potentials of Human Mesenchymal Stem Cells for Early Evaluation of Intact Cells. *PLoS ONE* 9(4): e93952. doi:10.1371/journal.pone.0093952

Editor: Christina Chan, Michigan State University, United States of America

Received: October 10, 2013; **Accepted:** March 11, 2014; **Published:** April 4, 2014

Copyright: © 2014 Sasaki et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors are grateful for financial support from the New Energy and Industrial Technology Development Organization (NEDO), which provided the Grant for Industrial Technology Research (Financial Support to Young Researchers, 09C46036a); the Japan Society for the Promotion of Science (JSPS) no. 25-8678; the Hori Sciences and Arts Foundation; and the Program for Leading Graduate Schools "Integrative Graduate Education and Research in Green Natural Sciences" from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors declare there are no competing interests except for the patents and employment stated below. Author RK received a research grant from Nikon Corporation for collaborative research titled "Development of quality assurance technology using cellular images" (18000 US dollars/year). Also, RK is one of the main investigators of 12 patents (pending) related to basic technologies for "algorithms of cellular image processing". Author YK is the manager of the Business Planning Department, Instrument Company, Nikon Corporation. He is a member of a collaborative project with Nagoya University and is also one of the inventors of 12 related patents. Author HS is one of the main investigators of 2 patents in 12 related patents. Related patents pending: WO 2010/098105, PCT/JP2010/005119, 2010-100162, 2010-100163, 2010-100164, 2010-100165, 2010-100166, 2010-100167, 2011-042210, 2011-096258, 2011-501512, US-2012-0092478-A1. However above all, the authors declare that this does not alter their adherence to PLOS ONE policies on sharing data and materials.

* E-mail: kato-r@ps.nagoya-u.ac.jp

Introduction

The application of recent advances in cell technologies in regenerative medicine holds great promise for revolutionizing conventional medical therapies [1]. However, the lack of assessment technology for quantitatively evaluating cell quality, in particular for revealing both the current properties and the future potentials of intact cells, is a technical obstacle to the development of quality-assured cellular products for medical use [2,3]. Conventional methods for cellular assessment using

standard techniques of molecular biology are incompatible with satisfying clinical requirements, because these methods damage cultured cells. As a result, manual microscopic monitoring, the basic and the most traditional scheme for maintaining cells, is still the most practical quality-control method for facilities that distribute regenerative cell therapies [4–6].

Technological advances in optical systems and image-processing technologies have changed the status of image-based data from an art, available only to experts, to a technique that can be used to

generate unbiased data. Many high-content image-analysis methodologies based on imaging and image-processing technologies, especially those focused on fluorescently labeled images, have contributed to advances in drug discoveries [7–10]. In the field of applied cell therapy, several reports have indicated that cellular morphological information, combined with sophisticated computational modeling approaches, can serve as a descriptive indicator in evaluations of stem cells [11–14]. However, to fulfill the clinical requirements for producing intact cells for therapies, wider use of cell-morphology analysis methodologies that focus on label-free images should be encouraged.

In an effort to overcome these limitations of conventional methods using fluorescently labeled images, we previously performed a model case study of the label-free morphology-based prediction of the osteogenic differentiation potential of human bone marrow-derived stem cells (hBMSCs) [15,16], using a technique that combines an automatic cell monitoring system with effective computational modeling [17]. Statistically extracted features of cellular morphologies clearly indicated that their information content can satisfactorily train computational models, not only to quantitatively evaluate current cellular status, but also to quantitatively forecast their future status, i.e., their potentials. The greatest advantage of our proposed morphology-based cell quality assessment is its non-invasiveness. As a result of this feature, our method has benefits that cannot be achieved by conventional techniques for producing cells for clinical regenerative medicine: (1) elimination of risk factors, e.g., contamination and mishandling by the operator; (2) synchronic and flexible scheduling of culture and clinical operations, for the best timing of cellular activity; and (3) repeated assessment of the same sample, by multiple criteria and at multiple times, yielding data that better reflects the complex and dynamic features of the samples. Such intelligent control of culture processes is also a key technology for process automation [18].

In this work, we expanded our previous efforts to predict single-lineage differentiation potentials [17] by pursuing five important aims: (I) Confirmation of the robustness of our method for adapting to the practical cellular variation. In our earlier work, it was not clear whether our original methodology was applicable to wider ranges of cellular variations. To investigate this issue, our data were expanded to cover eight continuous passages, ranging from very recently derived cells to those that had completely lost their doubling potential. Since a computational modeling solution for adapting to cellular variations resulting from patient diversity was already proposed in our previous work [17], our experimental design in this work was focused on cellular variations affected by culture processes, because these are the most difficult aspect of stem cells to evaluate daily. (II) Investigation of the possibility of shifting the prediction timing to the very early stage. Our previous prediction required 2 weeks of image acquisition after the differentiation process began [17]. In this study, however, we investigated whether much earlier and shorter periods were possible. In this work, only four images, obtained from the same sample repeatedly with a 24-hour interval during the first 4 days of expansion before differentiation culture, were used in the predictions. (III) Multiplication of the variations of *in silico* predictions. Compared to the previous prediction scheme [17], which could predict osteogenic differentiation potential from the same image, in this study we attempted to predict four types of potentials (osteogenic/adipocyte/chondrocyte differentiation, and population doubling time (PDT)) from the same image. Such simultaneous prediction of multiple potentials for the same cells can be achieved by processing the same image data, although the predictions are performed by four types of differently trained

prediction models running in parallel. Thus, this is a trial of “overlapping” computational evaluation that can compensate for multiple immunohistochemical staining. (IV) Establishment of new conversion schemes of morphological feature usage that can achieve high predictive performance. Morphological features are the essential information generated from imaging data, and use of this information is critical in imaging-based applications. To date, however, there have been few comprehensive studies that compare the effects of different conversions of morphological features, especially in the context of label-free time-course imaging data. To reveal differences resulting from the use of various morphological features, we proposed six types of novel morphological feature conversion methods, and then compared their prediction performances in detail. To interpret the patterns of morphological features engaged in high-performance models in each differentiation lineage, we selected LASSO regression as a modeling method. (V) Quantitative comparison of morphology and gene expression in prior prediction of differentiation potential. Although morphological information has long been used as an indicator for cellular evaluation, it has remained unclear how descriptive such information really is. To quantitatively compare the performance of morphological and biological information, we directly compared the performances of predictive models using morphological features, gene expression, or both in predicting differentiation potentials from the undifferentiated state. This comparison provides a performance benchmark for our proposed morphology-based cellular potential prediction methodology, enabling complete, non-invasive, daily cellular evaluations that could support or complement evaluations that rely on conventional biomarkers.

Results

Construction of a dataset that relates hBMSC morphological information with differentiation potential, for the purpose of developing a model for early prediction using undifferentiated status images

To construct the morphology-based cell-quality prediction model, we first designed to prepare the dataset of hBMSCs images and their experimentally determined differentiation potential data. To assemble this dataset, three lots of hBMSCs were continuously cultured (8 passages) until their growth terminated. The range of cells was intended to mimic the wide variations in cell qualities of clinical hBMSCs. At each passage, each sample was divided into three groups: passage sample (SEED group), pre-differentiation sample (PRE group), and differentiation sample (DIFF group) (Fig. 1). Because the diversity of our cell samples was intended to mimic the clinical situation, in which a minimum cell yield is often required to meet the production criteria, the passage timing was controlled by confluency. Specifically, passage was performed when confluency exceeded 80%. Continuous passage was maintained using the SEED group. Meanwhile, the PRE group was subjected to phase-contrast microscopic image acquisition (4 days, 24-h intervals), and the DIFF samples were differentiated into three mesenchymal lineages (osteogenic, adipogenic, and chondrogenic). After long-term differentiation into the three lineages, cells were evaluated for their differentiation rate and PDT; these data were taken to represent the biological differentiation potentials. In the dataset, these potentials were linked to the morphological features measured from images in the PRE groups by machine learning using the LASSO model. Because we sought to investigate the possibility of extremely early prediction of stem cell differentiation potentials for clinical applications, we acquired our image data,

which we expected to contain predictive information, before the differentiation process began. Ultimately, the full hBMSC dataset contained 24 samples of cell variants (3 lots \times 8 passages [P2–P9]); 80 images (5 fields of view \times 4 wells \times 4 time points) from each PRE group; and 296 experimentally determined differentiation values ($= 2 \times [18 \text{ images} \times 8 \text{ passages}] + [1 \text{ image} \times 8 \text{ passages}]$) (Table S1).

By monitoring morphological changes in response to continuous passage, we observed a clear morphological transition from a spindle shape into a flat and polygonal shape, a typical indication of decay of differentiation potential (Fig. 2A). The morphological information was quantitatively extracted as morphological features, as described in the Materials and Methods section.

Values related to differentiation into the three mesenchymal lineages revealed that continuous passage severely reduced the differentiation potential of hBMSCs (Fig. 2B–D, Fig. 3A–C). However, the transition patterns of differentiation potentials for the three lineages varied in a complex manner. Potentials to differentiate into adipogenic (Fig. 3B) and chondrogenic (Fig. 3C) lineages dropped rapidly, but these potentials were not correlated with the osteogenic differentiation potential (Fig. 3A). There were

also variations in the changes in differentiation potentials that could be attributed primarily to differences among patients. Lot A retained its chondrogenic differentiation potential for a relatively long period, but suddenly lost it after P7 (Fig. 3C). In Lots B and C, osteogenic differentiation potential changed dramatically during continuous passage (Fig. 3A). In Lot C, adipogenic differentiation potential was sustained in any passages (Fig. 3B). Therefore, as a summary of Fig. 3B, it was realized that the tendency of differentiation potentials between patient cells can be drastically disturbed by the effect of culture process. This result indicates that simple categorization of cells by “patient information” is not effective in the clinical cell production processes, and their daily evaluation is essential.

PDT changed relatively slowly between P2 and P8 in Lots B and C (Fig. 3D). By contrast, in the case of Lot A, a rapid increase in PDT (i.e., reduction in growth rate) was observed starting at P6, indicating that this lot was sensitive to passage-related stresses (known as culture process–derived stress) triggered by both manipulations and *in vitro* culture conditions [19]. The irregular PDT increase in Lot A could be an indication of loss of

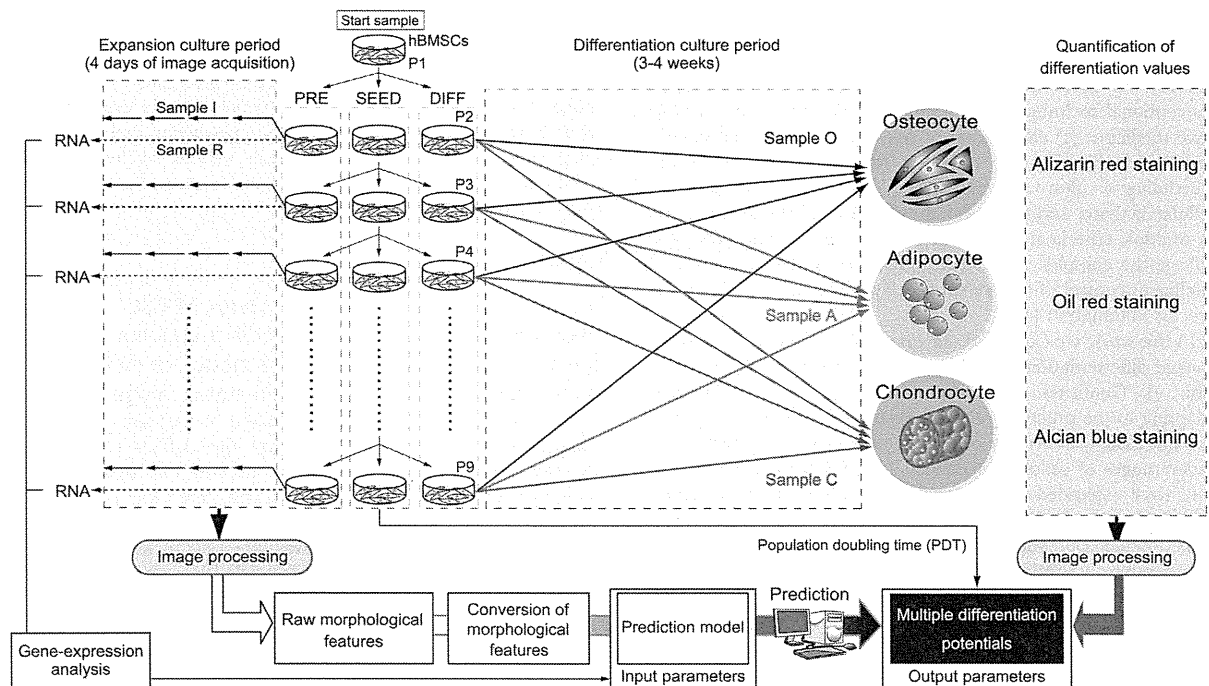


Figure 1. Schematic illustration of experimental setup for dataset construction for morphology-based prediction model construction. Figure S1 shows the illustration of usage of the objective morphology-based prediction model, and its major technological achievements using this dataset. The initial sample (P1) was divided into three separate culture samples (SEED, PRE, and DIFF) at each passage. SEED samples were mainly used for the continuous-passage culture until termination of growth (P9). From the cell yield at each passage of the SEED samples, population doubling time (PDT) was calculated, and taken as the experimentally determined potential. DIFF samples derived from each passage were divided into three differentiation cultures (samples O, A, and C for osteogenic, adipogenic, and chondrogenic differentiation, respectively) and grown under the indicated conditions for 3–4 weeks. The differentiation values of samples O, A, and C were experimentally quantified by individual staining potentials. The three types of differentiation potentials together with the population doubling time (population doubling time: PDT) were designated as “multiple differentiation potentials” of the hBMSCs. PRE samples consisted of sample I (for imaging) and sample R (for RNA extraction). From sample I in each passages, phase-contrast image were acquired at 24 h intervals over 4 days. Acquired images were then converted by image processing to obtain morphological features from every cell in all images (see also Fig. S2 and S3 for the details of image processing). Morphological features were statistically processed to yield transformed morphological features through data cleansing and statistical calculations, and the results were used as the input features. Sample R were subjected to total RNA extraction for gene-expression analysis. Either or both morphological features or/and gene-expression data were combined (input parameters), and arranged with the experimentally determined potentials of the hBMSCs (output parameters) to constitute training data for construction of prediction models. doi:10.1371/journal.pone.0093952.g001

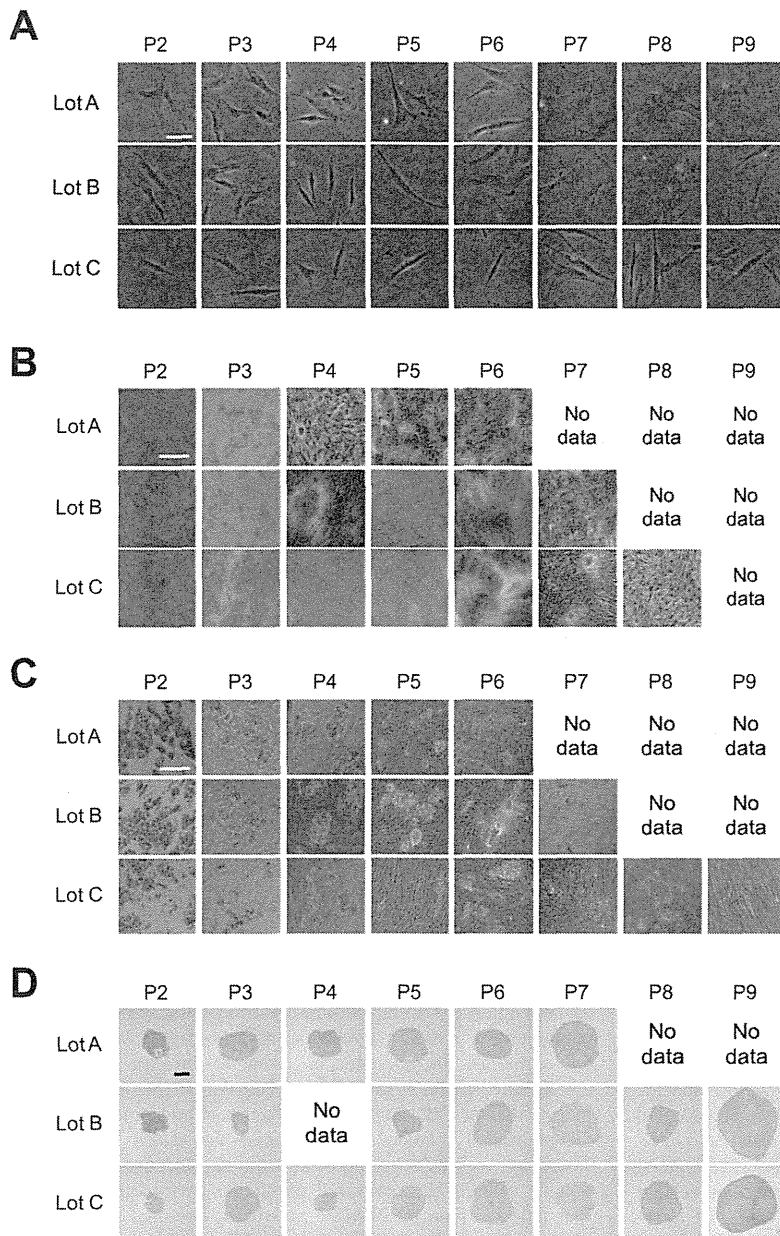


Figure 2. Representative morphological images of continuously passaged hBMSCs. Columns indicate passage numbers, indicated as P-number. Rows indicate hBMSC lot names. (A) Phase-contrast microscopic images ($10\times$) prior to differentiation culture (sample I). Scale bar, $50\ \mu\text{m}$. (low-resolution cellular images shown in Figure S4) (B) Alizarin red staining after 2 weeks of osteogenic differentiation culture (sample O). Scale bar, $200\ \mu\text{m}$. (C) Oil red staining after 3 weeks of adipogenic differentiation culture (sample A). Scale bar, $200\ \mu\text{m}$. (D) Alcian blue staining after 4 weeks of chondrogenic differentiation culture (sample C). Scale bar, $200\ \mu\text{m}$. From P7–P9, near the termination of growth, differentiation samples could not be prepared for (B) and (C) because of the lack of cell numbers. In (D), when the pellet sizes were smaller than $200\ \mu\text{m}$, we declined to produce specimens from the sample on the grounds that the differentiation culture had not been successful. doi:10.1371/journal.pone.0093952.g002

differentiation potential; however, such an indication does not explain the early change in adipogenic and chondrogenic differentiation potentials in Lots B and C. These results reveal that there are no simple correlations between passage number and transition patterns. If passage numbers or PDT do not reflect changes in differentiation potential, then this information would

never be sufficient to avoid a sudden quality collapse or insufficiency of cellular potential.

From gene-expression profiles of the earliest stage of expansion culture prior to differentiation, we found that most of the conventional differentiation markers did not exhibit clear synchronization with passage number (Fig. 4). Most of the

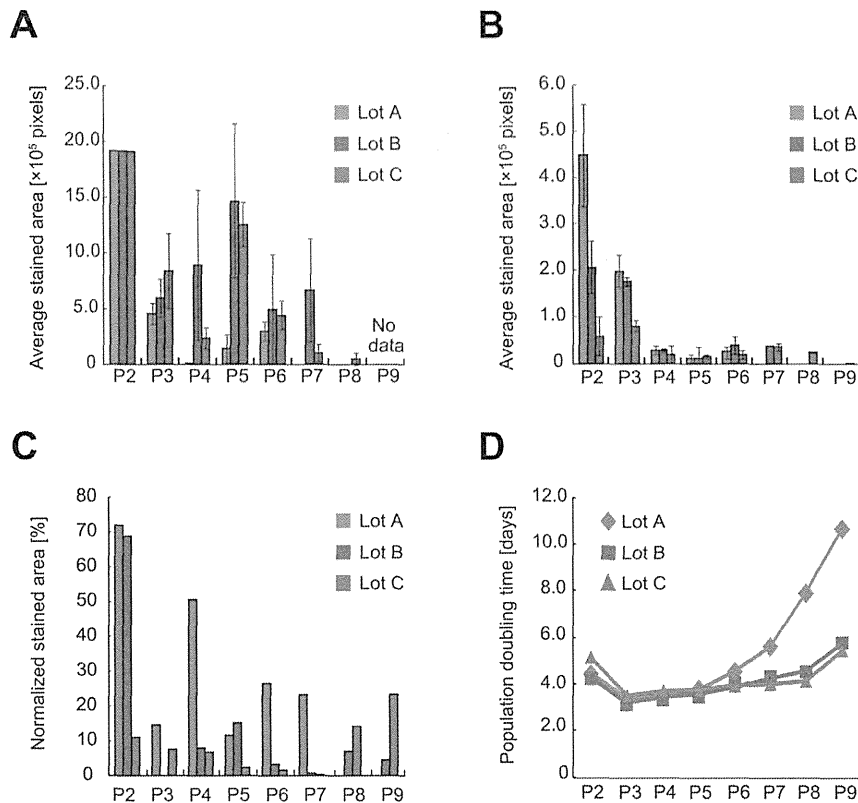


Figure 3. Quantified experimentally determined differentiation values and population doubling times of hBMSCs. Green bar, Lot A; blue bar, Lot B; pink bar, Lot C. Passage numbers are indicated as P2–P9. (A) Bar plots of average stained areas of Alizarin red–stained samples ($n=6$). (B) Bar plots of average stained areas of Oil red–stained samples ($n=6$). (C) Bar plots of stained areas in Alcian blue–stained samples ($n=1$), normalized by their pellet size. (D) Line plots of PDT. Green diamonds, Lot A; blue squares, Lot B; pink triangles, Lot C. Error bars indicate standard deviation (s.d.).
doi:10.1371/journal.pone.0093952.g003

clustered genes indicated the expression transition characteristic to certain cell lot. Therefore, there were several clusters, which partially showed synchronization to passage number in certain lot was considered to be more genes that reflect patient specific response to passages. However, in clustered genes which indicated clear synchronization with passage number (correlation coefficient >0.673) among all cell lots, either cellular senescence–related genes (*CDKN1A* [p21], *CDKN2A* [p16]) or cytoskeleton–related genes (*PTK2*, CD146 [*MCAM*], and CD49 [*ITGA1*]) were included. Since passage number significantly correlated with the decrease of differentiation potentials (Fig. 3), such passage number synchronizing gene expressions commonly observed in all cell lots were considered to be the “genetic signature of potential collapse”.

Comparison of performances of prediction models to achieve the most balanced performance

In our previous study, we found that time courses of morphological features of cultured hBMSCs were informative in the construction of computational models aimed at forecasting future osteogenic differentiation [17]. To evaluate the multiple potentials of hBMSCs in practice, our concept of prediction had to be expanded from single-lineage to multi-lineage differentiation while retaining the ability to adapt to wider cellular variations. However, we hypothesized that in order to predict multi-lineage

differentiation potential, informative morphological features and combinations thereof should be optimized for each type of differentiation. In addition, in order to increase the clinical applicability of this approach, our conceptual prediction models had to balance the accuracy with the effort (time, cost, and computational memory size) required for model preparation. Therefore, we have set our goal to define the effective construction scheme yielding the optimized prediction performance for each four different types of hBMSC potentials: potential I, osteogenic differentiation rate after 2 weeks of differentiation; potential II, adipogenic differentiation rate after 3 weeks of differentiation; potential III, chondrogenic differentiation rate after 4 weeks of differentiation; and potential IV, PDT of cells after the passages. Our objective prediction model with the newly developed techniques in this work is illustrated in Fig. S1.

To achieve the best prediction models for these objectives, we examined nine patterns (Models 1–9) of input data usages, which critically change the users’ efforts for data preparation. Model 1 was designed to be the negative control, and Models 2 and 3 were designed to compare gene expression–based predictions compared to morphology–based models. Models 4–9, consisting of five model patterns (M-patterns), were designed to compare morphological feature conversion methods by investigating the various conversion concepts and time-course data usage in morphological features (Fig. 5). The comparison of morphological features was

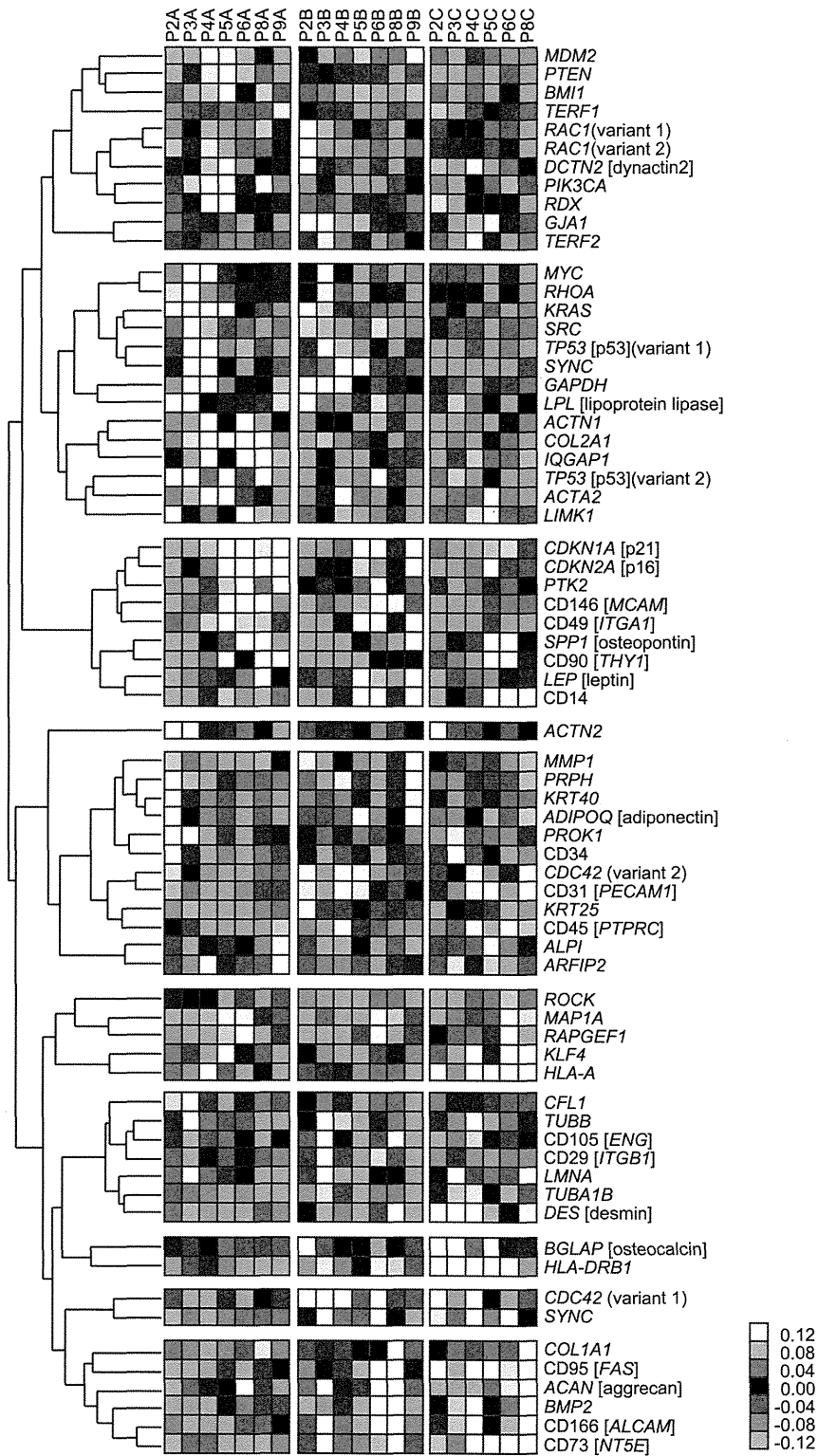


Figure 4. Heat map of gene-expression transitions and passage numbers. Genes were clustered by hierarchical clustering for indicating clusters that correlate to the passage number increases. The red boxed cluster is the cluster of genes that correlate to passage number within all cell lots, indicating non-patient-specific passage increases. The relationship between colors and normalized values of gene expression is illustrated in the explanatory heat map at lower right.
doi:10.1371/journal.pone.0093952.g004

deepened by balancing accuracy vs. feasibility of usage in the clinic. Therefore, M-patterns were numbered in the order of higher cost performance, considering the time, cost, and computational memory size involved in the model-construction process. All model performances were compared by two criteria, “scaled error rate” and “correlation coefficient”. The scaled error rate indicates the median value of prediction errors among all the samples, normalized by the actual experimental values. Therefore, low scaled error rate indicates that prediction values are relevant to the experimental values. However, usage of the scaled error rate alone is vulnerable to accidental prediction noises. Therefore, we introduced the second criterion, correlation coefficient, which evaluates the combined correlations of experimentally determined values and predicted values among all samples. These criteria are complementary: scaled error reflects differences between plots, however discards information about overall plot accuracy, whereas correlation coefficient reflects the overall similarity of measurements and predictions, however is sensitive to outliers. The combination of low scaled error rate and high correlation coefficient indicates stable performance of a given model.

Figure 6 depicts all prediction results. The data indicate that by using only the prior morphologies before the differentiation process, future collapses in all of differentiation potentials (potential I–IV) under continuous passage stresses can be predicted in advance. Comparisons of the transition patterns of the experimentally determined and predicted values (blue line plots and red line plots, respectively, depicted in Fig. 6) revealed that all cellular properties were predicted with reasonable accuracy. Furthermore, in contrast to our previous study that used all morphological data from 14 days of differentiation culture period [17], the predictive performance was enhanced in this study using morphological data collected only from the first 4 days before the differentiation.

For prediction of potential I (osteogenic differentiation rate), the best prediction accuracy was achieved by Model 3, which utilizes both morphological features (M-pattern 1) and gene-expression profiles. Compared to the scaled error rate of the NULL model (Model 1), the performance of Model 3 can be expressed as 2.6-fold more accurate. In the sense of cost-efficiency of model construction, Model 8, which utilizes only the morphological features from 24 h, had a reasonably high predictive performance (scaled error rate = 0.338).

For predicting potential II (adipogenic differentiation rate), morphology-based models such as Models 4, 6, 7, and 9 yielded extremely high predictive performance. Model 4 achieved the best accuracy, and Model 7 was the best model at the lowest cost (scaled error rate = 0.954).

For predicting potential III (chondrogenic differentiation rate), most of the models could not significantly outperform the NULL model. However, Model 9 had fairly accurate predictive performance.

For predicting potential IV (PDT of cells after repeated passages), most of the models had very high predictive accuracies (scaled error rate < 0.09). The best performance was achieved by Model 4, which utilizes only morphological features.

Although morphology-based prediction models (Models 4–9; M-Patterns 1–5) had consistently high overall performances in predicting various potentials, the use of direct biological information (i.e., gene-expression information including conventional differentiation markers) did not dramatically improve the predictive performance (Model 2 in Fig. 6). From the interpretation of parameter usages in LASSO models (Table S4–S7), we found that cytoskeleton-related genes were more frequently involved than differentiation markers in the prediction models (Models 2 and 3).

This result is a biological confirmation that morphological genes are more informative than our selected subset of differentiation marker genes for prediction of differentiation potential, and explains the high performances of models that use only morphological data (Models 4–9).

Discussion

To replace human estimations of cell quality in the production of cells for cell-based therapies, we examined the performances of machine-learning models in predicting the quantitative rates of multi-lineage differentiation after long-term differentiation, using data from undifferentiated label-free images of hBMSCs. The novel advancing technological points achieved in this work are illustrated in Fig. S1. From images collected during the first 4 days of expansion culture before differentiation, the morphological features of each cell in the images were individually measured and converted into various morphological metrics that represented the statistical morphological profiles of the group of cells. These features were then used to train computational models that forecast the experimental results collected 2–4 weeks after the differentiation. Advancing from our previous success in predicting the single-lineage differentiation potentials of hBMSCs [17], here we showed that the best predictive results for all differentiation potentials (i.e., the differentiation rates into three lineages and their PDTs) can be obtained at the same time, even in the early stage before differentiation, using selected morphological features. In these comparisons, we addressed three technical questions, with the aim of identifying the most practical scheme for obtaining such cell-quality prediction models in clinical facilities. First, can morphology-based prediction methods be expanded to the prediction of multiple differentiation potentials? Second, is morphological information (i.e., indirect phenotypic signals) of greater use than gene-expression information (i.e., direct biological signals) in predicting the qualities of hBMSCs? Third, how far can we optimize model performance by selecting the appropriate conversion and combination of information from the time-course morphological features?

To our great surprise, considering the current lack of comparable evaluation methods, most of the examined prediction models using only morphological features showed practically useful performance in multiple predictions (Fig. 5 and 6). Even with the Model 9 (M-pattern 6, using morphological features obtained only from the first day of expansion culture), the multi-lineage potential prediction was available. Practically, potential II (adipogenic differentiation rate after 3 weeks) can be predicted with high accuracy using only morphological data from the first 4 days of culture. Both potentials I and III (osteogenic and chondrogenic differentiation rates) could also be predicted with reasonable accuracy from the early morphological data. In addition to differentiation rates, future PDT following repeated passages can also be predicted with high accuracy using only morphological features. These results strongly indicate that it will be possible to develop practical methods for cell assessment that are multiple, rapid, cheap, non-invasive, and significantly more effective than conventional staining-based assessment techniques. Our models’ performance indicate that such novel predictive methods will enjoy several advantages: (1) non-invasiveness, i.e., avoiding damage to patients’ cells; (2) synchronism, repeated quality evaluation throughout the culture period for all patients; and (3) multivalent consideration of the same sample, i.e., multiple quality assessments can be performed with the same sample, which is not possible when using data obtained by destructive methods such as fluorescently labeled imaging analysis.

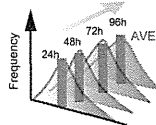
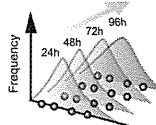

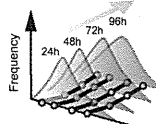
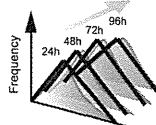

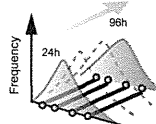
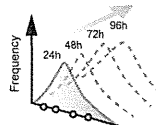
M-pattern	Parameter conversion profile							Parameter usage		
	Statistical values from cellular objects	Linkage of parameters through time points	Time points [h]				Number of parameters			
			24	48	72	96	Morphological parameters		Counts	Total
1	Average	Non-linked	•	•	•	•	36	4	40	
2	Quantile points	Non-linked	•	•	•	•	180	4	184	
3		Linked					135	3	138	
4	Distribution representative (median, interquartile range, robust skewness, robust kurtosis)	Non-linked	•	•	•	•	144	4	148	
5	Quantile points	Linked					45	1	46	
6		Non-linked	•				45	1	46	

Figure 5. Schematic illustration matrix of prediction feature data profile and usage concepts of prediction models. Six types of morphological feature conversion methods are proposed as M-patterns. Briefly, M-patterns are numbered in order of the amount of efforts required to prepare for model construction. M-patterns 1–4 require four images at 24-hour intervals; M-pattern 5 requires two images each on days 1 and 4; and M-pattern 6 requires only one image on the first day. For parameters described as “linked”, each morphological feature is not only used as the data for each time point, but this information is also converted into the changing ratio between time points. For “non-linked” parameters, morphological features are used as they are. Averages, quintile points, and groups of distribution representatives were compared to find the best statistical parameter to represent the morphological features measured in all individual cells in an image. Therefore, M-patterns 1–4 were designed to increase the amount of information about cellular distribution for incrementing the heterogeneity of cells. doi:10.1371/journal.pone.0093952.g005

The quantitative predictions made possible by these methods will permit prior evaluation of cellular fate, which will in turn facilitate scheduling of cell-therapy operations in the clinic. As shown in Fig. 3, most of the transition events in hBMSC potentials were abrupt, and would be nearly impossible to estimate the future linearly from the present result plots. Therefore, conventional cell-assessment techniques could never outperform quantitative prediction methods for hBMSC quality assessment. Our results thus provide a successful example of the use of machine-learning models to model biological information and generate output that can overcome a major practical problem in clinical cell therapy.

Taken together with the non-linear correlation of conventional marker gene-expression levels with passage numbers (Fig. 4) and

the predictive performance of models (Models 2 and 3 in Fig. 6), we concluded that morphological data from the early stage of culture are more useful than measurements of conventional markers in forecasting future quality disruptions. In some cases, gene-expression measurement enhanced morphological predictions, when an early gene marker such as *SPPI* [osteopontin] occasionally function as extreme early osteogenesis predictor (Model 2, Potential I prediction in Fig. 6). However, differentiation gene markers are not always promising to function as extreme early predictor in the undifferentiation stage. By introducing LASSO modeling into this work, the combinational effects of parameters can be interpreted in our models (Fig. 6). In particular, by interpreting the parameter usages chosen through automatic